2Б

УДК 519.854:004.8

Н.В. Белоус, И.В. Куцевич
Харьковский национальный университет радиоэлектроники, г. Харьков, Украина
belous@kture.kharkov.ua

Дифференциальное оценивание знаний при дистанционном тестировании

В статье исследованы основные подходы к созданию тестов, приведена классификация форм тестовых заданий. В работе авторами предлагаются принципы построения дифференциальной модели процесса тестирования, поддерживающей индивидуальность оценивания заданий разных форм, а также приводится технология оценивания вероятности угадывания правильных ответов для каждой из форм тестовых заданий.

Особое внимание в системе дистанционного обучения (ДО) уделяется контролю качества усвоения теоретического материала и тестированию приобретенных знаний и практических навыков обучаемого. Решение проблемы контроля качества ДО и соответствие образовательным стандартам имеет принципиальное значение для успеха всей системы ДО. Кроме того, при удаленном тестировании и контроле должны решаться проблемы индивидуализации процесса обучения и контроля знаний. Для решения этой проблемы разработано адаптивное тестирование знаний. Адаптивный тест представляется как вариант автоматизированной системы тестирования, в которой заранее известны параметры трудности и дифференцирующей способности каждого задания. Эта система создана в виде компьютерного банка заданий, упорядоченных в соответствии с интересующими характеристиками заданий. Основной из задач разработки адаптивного теста является построение системы оценивания адаптивного тестирования.

В настоящий момент в зарубежной педагогической практике существуют несколько подходов к созданию тестов, отвечающих научно-обоснованным критериям качества. Первый подход базируется на применении стандартных для математической статистики корреляционных и факторных методов анализа. Второй современный подход зарождался в рамках теории латентно-структурного анализа, основная цель которого – оценить скрытые (латентные) параметры испытуемых. Пионерской работой в области второго подхода является труд датского математика Г. Раша [1], вышедший в 1960 г. Именно он дал толчок интенсивному развитию за рубежом серьёзной теоретической базы тестирования, вылившейся в современную теорию тестирования – Item Response Theory (IRT). Первые сведения по этой теории на русском языке появились только в 1995 г. в работе Р. Хамблетона [2].

Основная цель состоит в разработке математической модели процесса тестирования, параметрами которой, подлежащими определению, служат различные характеристики участников тестирования и самого теста. В настоящее время IRT представляет собой довольно обширную и, к сожалению, непростую теорию, активно использующую разнообразный арсенал математических методов [3]. Поэтому зачастую целесообразно бывает использовать сочетание двух упомянутых теорий. С помощью статистических методов классической теории проводится первичный анализ качественных характеристик полученного варианта теста, а с помощью IRT, более углублённый анализ характеристик

тестовых заданий. Наиболее эффективным использование IRT становится при сформированном банке тестовых заданий, что обеспечивает возможность автоматической генерации тестов с определёнными наперёд заданными и научно-обоснованными характеристиками точности и надёжности.

Основными недостатками рассмотренных моделей является использование дихотомической и политомической шкал оценивания знаний, а также относительная сложность реализации применяемых математических методов при построении систем адаптивного дистанционного тестирования знаний.

Целью статьи является разработка дифференциальной модели процесса адаптивного тестирования знаний при прохождении дистанционного обучения, учитывающей уровень сложности заданий, вероятность угадывания правильных ответов, а также учитывающей оценивание заданий различных форм по непрерывной шкале оценивания.

1. Построение дифференциальной модели процесса тестирования

При прохождении адаптивного теста в каждый момент времени обучаемые выполняют некоторые задания. По результатам выполнения каждого из заданий выдвигаются условия для выбора следующего задания из базы знаний.

Исходя из сказанного выше, построим дифференциальную модель процесса оценивания знаний с помощью тестирования. Дифференциальная модель представляет собой дифференциальное уравнение, полученное в результате исследования какого-либо реального явления или процесса [4], в данном случае, процесса оценивания теста.

Итак, пусть некоторый обучаемый проходит адаптивное тестирование. Ему предлагается выполнить цепочку из N заданий, причем результат выполнения теста R в каждый момент времени (при выполнении каждого последующего задания) зависит от результатов, полученных при выполнении предыдущих заданий. Таким образом, результат выполнения теста изменяется непрерывно и, более того, он дифференцируем как функция, зависящая от количества выполненных заданий. Конечно, это утверждение является упрощением реальной ситуации, поскольку R – целое число.

Для построения модели проведем анализ процесса адаптивного тестирования. Рассмотрим процесс прохождения тестирования и выделим основные параметры, влияющие на результат прохождения тестирования.

1. При прохождении адаптивного тестирования каждому из заданий присваивается уровень сложности z_i ($z_i = \overline{1, Z}$). В связи с этим, чем больше вероятность выполнения задания правильно, тем ниже уровень сложности задания.

2. Для получения объективной оценки знаний при прохождении тестирования рекомендуется использовать непрерывную шкалу оценивания. При этом за выполнение каждого из заданий теста обучаемый получает некоторый коэффициент $r_i \in [0;1]$, причем этот коэффициент определяется индивидуально для каждой из форм тестовых заданий.

3. Одной из основных проблем тестирования является вероятность угадывания правильных вариантов ответа. Следовательно, в дифференциальную модель процесса проведения адаптивного тестирования введем параметр w_i, соответствующий вероятности угадывания правильного варианта ответа для каждой их форм тестовых заданий.

4. Для получения результата в заданной шкале оценивания знаний введем специальный параметр B – балльность системы. Рассмотрим алгоритм перевода результата тестирования в произвольную шкалу оценивания знаний. Для корректного перевода результата в любую систему оценивания знаний необходимо всем значимым результатам оценивания в порядке возрастания поочередно присвоить коэффициенты балльности, начиная с 1. Максимальное значение коэффициента балльности и будет соответствовать параметру B.

При учете введенных параметров, влияющих на результат прохождения теста, динамика изменения результата выполнения теста при выполнении каждого из заданий может быть описана формулой 1.

$$\frac{dR}{dN} = \left(r_i \cdot z_i + \frac{dR}{d(N-1)} \cdot \frac{(1-w_i)(\dot{Z}-z_i)}{B} \right) \cdot \frac{B}{\dot{Z}}, \qquad (1)$$

где $\frac{dR}{dN}$ – дифференциальное оценивание результата прохождения теста в данный момент времени;

$\frac{dR}{d(N-1)}$ – дифференциальное оценивание результата прохождения теста в предыдущий момент времени (при выполнении ($N-1$) задания);

N – количество тестовых заданий, выполненных обучаемым в данный момент времени;

r_i – коэффициент оценивания заданий разных форм. Непрерывная величина изменяется в диапазоне [0,1];

z_i – уровень сложности i-го задания;

w_i – вероятность угадывания правильного ответа;

B – балльность системы;

$\dot{Z}$ – суммарная сложность теста. Параметр $\dot{Z}$ является накопительным и определяется по формуле 2.

$$\dot{Z} = \sum_{i=1}^{N_t} z_i , \qquad (2)$$

где i – номер задания;

N_t – номер текущего выполняемого задания.

Задача состоит в том, чтобы указать соответствующие формулы для введенных выше величин коэффициента оценивания заданий разных форм r_i и вероятности угадывания правильного ответа w_i.

2. Классификация форм тестовых заданий

В настоящее время существует несколько классификаций форм тестовых заданий [5-7]. Авторами предлагается при построении теста использовать формы тестовых заданий, приведенных в классификационной схеме, представленной на рис. 1 [8].

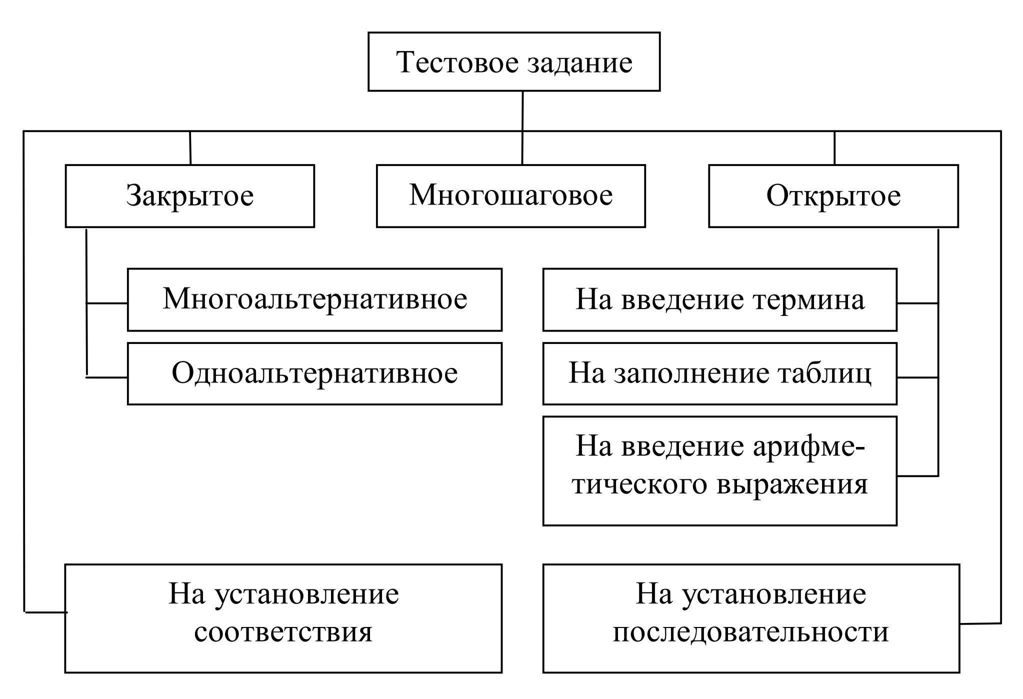


Рисунок 1 – Классификационная схема тестовых заданий

3. Определение коэффициентов оценивания заданий разных форм

Для того, чтобы объективно оценить знания по тестовым заданиям разных форм предлагается использовать для каждой из них свою методику расчета оценки. Определим коэффициент оценивания для ответов на тестовые задания разных форм в отдельности.

3.1. Одноальтернативные тестовые задания

Для оценивания одноальтернативного тестового задания достаточно применение известной и широко используемой дихотомической шкалы оценивания, где 1 соответствует правильному ответу, 0 – неправильному. Следовательно, в этом случае коэффициент r_i принимает одно из значений 0 или 1.

3.2. Многоальтернативные тестовые задания

При оценивании многоальтернативных заданий этой шкалы недостаточно, так как обучаемый может дать неполный ответ, либо один из выбранных вариантов ответа будет неточен. Такие ответы нельзя оценивать так же, как и вопросы, в которых был выбран полностью неправильный ответ. В случае многоальтернативного тестового задания необходимо учитывать не только правильность ответа на задание в целом, но и количество вариантов ответа, выбранных обучаемым. В данном случае коэффициент r_i предлагается рассчитывать по формуле 3.

$$r_i = \frac{Q_{2_s}}{(Q_{1_i} + Q_{3_i})}\ , \qquad (3)$$

где Q_1 – множество всех правильных вариантов ответа в задании,
Q_2 – количество правильных вариантов ответа, выбранных обучаемым,
Q_3 – количество неправильных вариантов ответа, выбранных обучаемым.

3.3. Задания на установление соответствия

При ответе на задание на установление соответствия каждую пару ответов можно рассматривать как отдельный вариант ответа и при выставлении результата выполнения задания следует учитывать, сколько пар было выбрано верно. Следовательно, коэффициент r_i необходимо рассчитывать по формуле (4).

$$r_i = \frac{Q_{2_i}}{Q_{1_i}} , \quad (4)$$

где Q_1 – количество пар для сопоставления;
Q_2 – количество верно составленных пар.

3.4. Задания на установление последовательности

При оценивании заданий на установление правильной последовательности возможен только один заведомо правильный ответ. Следовательно, для оценивания данной формы тестовых заданий достаточно использовать дихотомическую шкалу оценивания, коэффициент r_i принимает значение 0 или 1.

3.5. Задания на заполнение таблиц

Одной из форм открытых тестовых заданий является заполнение таблиц. Каждая ячейка таблицы является отдельным вариантом ответа и, если одна из ячеек заполнена неправильно, такой ответ нельзя засчитывать как полностью неправильный (введение одного неверного значения в ячейку может быть лишь механическая ошибка, и поэтому оно должно не полностью обнулить результат выполнения работы, а лишь снизить результат выполнения задания). Для определения коэффициента оценивания заданий на заполнения таблиц рекомендуется использовать показательную функциональную зависимость, представленную формулой (5).

$$r_{z_i} = 2^{\frac{Q_{2_i}}{Q_{1_i}}} - 1, \quad (5)$$

где Q_1 – количество ячеек, которые предлагается заполнить тестируемому;
Q_2 – количество ячеек, которое тестируемый заполнил правильно.

3.6. Многошаговые тестовые задания

Многошаговые тестовые задания состоят из набора заданий (набора шагов), решаемых последовательно, когда переход к следующему шагу задания осуществляется только после правильного ответа на предыдущий шаг. Это дает возможность

обучаемому анализировать не только задание в целом, но и разбираться в каждой составляющей задания. Благодаря этому обучаемый сразу может увидеть, где им допущена ошибка и в дальнейших шагах получить правильные исходные данные, то есть ошибки в заданиях не будут накапливаться. Многошаговое задание считается пройденным, если на каждом его шаге получен правильный ответ. Для объективного оценивания ответа и глубины знаний обучаемого используется счетчик допускаемых ошибок, количество которых учитывается при выставлении оценки.

Многошаговое задание считается пройденным, если на любом его шаге получен правильный ответ. Для объективного оценивания ответа и глубины знаний обучаемого используется счетчик допускаемых ошибок, количество которых учитывается при выставлении оценки. Однако встает вопрос об оценивании таких вопросов. Для них простая дихотомическая шкала не подходит. Рекомендуемая формула для вычисления коэффициента правильности для многошагового тестового задания имеет вид (6).

$$r_i = \frac{\sum_{i=1}^{n} \frac{1}{m_i + 1}}{n}, \qquad (6)$$

где i – номер шага,

m_i – количество ошибок, допущенных на i-м шаге,

n – количество шагов.

Формула (6) справедлива для многошаговых тестовых заданий, в которых на каждом шаге используются одноальтернативные задания или задания на установление правильной последовательности. В случае использования на каком-либо из шагов тестового задания на соответствие или многоальтернативного тестового задания целесообразно использовать формулы (7) и (8) соответственно.

$$r_i = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{m_i}\sum_{j=1}^{m_i}\frac{Q_{1_{i_j}}}{Q_{2_{i_j}}}\right), \qquad (7)$$

где j – номер попытки прохождения шага, если на нем была допущена ошибка;

$Q_{2_{i_j}}$ – количество пар для составления на i-м шаге при j-й попытке;

$Q_{1_{i_j}}$ – количество верно составленных пар на i-м шаге при j-й попытке.

$$r_{z_i} = \frac{1}{n}\sum_{i=1}^{n}\frac{m_i \cdot Q_{2_i}}{\sum_{j=1}^{m_i}(Q_{1_{i_j}} + Q_{3_{i_j}})}, \ , \qquad (8)$$

где $Q_{1_{i_j}}$ – количество правильных вариантов ответов i-м на шаге при j-й попытке;

Q_{2_i} – количество правильных ответов, выбранных тестируемым на i-м шаге;

$Q_{3_{i_j}}$ – количество неправильных ответов, выбранных тестируемым на i-м шаге при j-й попытке.

4. Оценивание вероятности угадывания правильных ответов для каждой из форм тестовых заданий

Использование выборочных ответов сопряжено с возможностью неадекватной оценки знаний обучаемого в случае случайного ввода правильного ответа. Оценим эту вероятность для ответов на тестовые задания разных типов.

4.1. Закрытое тестовое задание

Для ответа на тестовые задания закрытого типа рассмотрим в отдельности ответ на одноальтернативные и многоальтернативные тестовые задания.

Одноальтернативное тестовое задание.

Одноальтернативные тестовые задания представляют собой задания, в которых N вариантов, ответа один из которых правильный. Вероятность угадывания правильного ответа w_1 в данном случае определяется по формуле 9.

$$w_1 = \frac{1}{N}. \qquad (9)$$

Многоальтернативное тестовое задание.

При выполнении многоальтернативных тестовых заданий обучаемому предлагается из N вариантов ответа выбрать k правильных ($0 \le k \le N$). Вероятность ввода каждого ответа для выборки длиной k определяется как $w_2 = \frac{1}{M}$, где M – общее количество возможных ответов. В случае, если заведомо известно количество правильных вариантов ответа, значение M рассчитывается по формуле 10.

$$M = C_N^k = \frac{N!}{(N-k)!k!}. \qquad (10)$$

Общее число возможных ответов M на многоальтернативное задание при заранее неизвестном числе k равно количеству всех возможных вариантов выбора ответа. Следовательно,

$$M = \sum_{k=0}^{N} C_N^k = \sum_{k=0}^{N} \frac{N!}{(N-k)!k!}.$$

Из комбинаторики известно [9], что $\sum_{k=0}^{N} C_N^k = 2^N$. Значение k неизвестно обучаемому, следовательно, можно считать все значения k равновероятными. Только один из возможных вариантов ответа является правильным. Таким образом, вероятность угадывания правильного ответа определяется по формуле 11.

$$w_2 = \frac{1}{2^N}. \qquad (11)$$

4.2. Тестовое задание на установление правильной последовательности

При составлении тестового задания на установление последовательности возможны 2 варианта представления списка элементов последовательности. В первом случае обучаемому необходимо составить последовательность, включив все элементы последовательности. Во втором случае предлагается выбрать последовательность элементов, включив в нее только необходимые элементы. Рассмотрим оба случая.

– Задание, в котором все N элементов последовательности входят в ответ.

Для заданий, в которых необходимо составить последовательность из всех предложенных элементов, задача определения количества возможных комбинаций ответа сводится к нахождению числа перестановок элементов последовательности. Таким образом, количество M возможных ответов определяется по формуле $M = N!$. Таким образом, вероятность угадывания правильного ответа для списка определяется по формуле 12.

$$w_3 = \frac{1}{N!}. \qquad (12)$$

– Задание, в котором из N предложенных элементов k входит в правильную последовательность.

Для рассматриваемого случая минимальное значение k равно 2, так как последовательность меньшей длины не имеет смысла. Количество M возможных вариантов ответа можно определить как размещение из N по k. Следует также учесть, что значение k изменяется от 2 до N. Таким образом, количество M различных вариантов перестановок определяется по формуле 13.

$$M = \sum_{k=2}^{N} A_N^k = \sum_{k=2}^{N} \frac{N!}{(N-k)!}. \qquad (13)$$

Правильным является единственный вариант, поэтому вероятность случайного ввода правильного ответа w_4 определяется по формуле 14.

$$w_4 = \frac{1}{M} = \frac{1}{\sum_{k=2}^{N} \frac{N!}{(N-k)!}}. \qquad (14)$$

4.3. Тестовое задание на установление соответствия

При выполнении тестовых заданий на установление соответствия возможно представление ответа одним из следующих способов:

1. Имеется 2 списка по m элементов. Необходимо составить m пар.

2. Имеется 2 списка по m и n элементов соответственно ($m < n$). Необходимо составить m пар.

3. Имеется 2 списка по m и n элементов соответственно ($m \leq n$). Необходимо составить l пар ($l < m$).

Рассмотрим каждый из способов в отдельности.

1. При установлении соответствия между элементами двух множеств с одинаковым количеством элементов общее количество возможных вариантов введения

ответа $M = (m!)^2$. Следовательно, в этом случае вероятность случайного ввода правильного ответа определяется по формуле 15.

$$w_5 = \frac{1}{(m!)^2}. \tag{15}$$

2. В случае, когда необходимо установить соответствие между m парами, причем количество элементов в одном из множеств n ($m \leq n$), количество возможных вариантов ответа возрастет и будет определятся по формуле 16.

$$M = P_m A_n^m = \frac{m!n!}{(n-m)!}. \tag{16}$$

Вероятность угадывания ответа в таком случае определяется по формуле 17.

$$w_6 = \frac{(n-m)!}{n!m!}. \tag{17}$$

3. Для получения задания на установление соответствия с наименьшей вероятностью угадывания рекомендуется добавлять «лишние» элементы в оба множества элементов. Пусть имеется 2 множества, состоящие из m и n элементов соответственно. При установлении составления для l пар ($l<m$, $l<n$) общее количество возможных ответов можно определить по формуле 18.

$$M = A_m^l A_n^l = \frac{m!n!}{(m-l)!(n-l)!}. \tag{18}$$

Вероятность угадывания ответа в случае, когда в задании на установление соответствия в оба множества элементов добавлены «лишние» элементы, определяется по формуле 19.

$$w_7 = \frac{(m-l)!(n-l)!}{m!n!}. \tag{19}$$

4.4. Открытое тестовое задание

При выполнении открытых тестовых заданий обучаемому не предлагаются варианты ответа, поэтому вероятность угадывания правильного ответа стремится к нулю.

4.5. Многошаговое тестовое задание

При оценивании вероятности угадывания правильного ответа во время выполнения многошагового тестового задания первоначально необходимо оценить вероятность угадывания правильного ответа w_{s_i} на каждом из l шагов в отдельности. Общий результат в данном случае определяется по формуле 20.

$$w_8 = \prod_{i=1}^{l} w_{s_i}. \tag{20}$$

4.6. Сравнительный анализ форм тестовых заданий учитывая вероятность угадывания правильного ответа

Рассмотрим вероятности случайного ввода правильного ответа в зависимости от количества вариантов ответа для тестовых вопросов с одним списком ответов. Для наглядности сравнения составим табл. 1, содержащую значения вероятности угадывания для тестовых заданий с количеством ответов от 3 до 7.

Таблица 1 – Вероятности угадывания правильного ответа для тестовых заданий с одним списком ответов

Количество заданий, N	Форма тестового задания			
	Одноальтернативное тестовое задание	Многоальтернативное тестовое задание	Тестовое задание на установление последовательности	
			Включающее все элементы	Включающее не все элементы
	w_1	w_2	w_3	w_4
3	0,33333	0,12500	0,16667	0,08333
4	0,25000	0,06250	0,04167	0,01667
5	0,20000	0,03125	0,00833	0,00278
6	0,16667	0,01563	0,00139	0,00051
7	0,14286	0,00781	0,00019	0,00007

На основании анализа полученных данных можно сделать следующие выводы:
– многоальтернативное тестовое задание значительно предпочтительнее одноальтернативного;
– тестовое задание на установление правильной последовательности, естественно, имеет меньшую вероятность случайного ввода правильного ответа, чем закрытое тестовое задание;
– при наличии «лишних» вариантов в ответе на задание на установление правильной последовательности его оценка улучшается еще приблизительно в 2 – 3 раза.

В свою очередь, следует отметить, что тестовое задание на установление соответствия имеет еще меньшую вероятность угадывания правильного ответа при соизмеримом количестве вариантов ответа с рассмотренными выше формами тестовых заданий. В свою очередь, для уменьшения вероятности угадывания рекомендуется в одно из множеств, между которыми устанавливается соответствие (либо в оба) вводить «лишние» элементы.

Выводы

Предложенная модель дифференциального оценивания знаний универсальна для применения в различных системах оценивания, а также проста для использования при реализации систем адаптивного дистанционного тестирования знаний. Поскольку при проведении тестирования существует возможность неадекватной оценки знаний обучаемого в случае случайного ввода правильного ответа, в работе проведена оценка предложенных форм тестовых заданий с точки зрения величины такой вероятности.

Полученные результаты планируется использовать при построении модифицированной системы проведения дистанционного адаптивного тестирования знаний в системе дистанционного образования Харьковского национального университета радиоэлектроники по курсу «Основы дискретной математики».

Литература

1. Rash G. Probabilistick Models for Some Intelligence and Attainment Tests. – Copenhagen (Denmark): Danish Institute for Educational Research, 1960.
2. Челышкова М.Б. Разработка педагогических тестов на основе современных математических моделей. – М., 1995. – С. 31.
3. Win J.van der Linden, R.K. Hambleton. Handbook of Modern Item Response Theory. – New Yourk: Springer-Verlag, 1997. – P. 510.
4. Амелькин В.В. Дифференциальные уравнения в приложениях. – М.: Наука. Главная редакция физико-математической литературы, 1987. – 160 с.
5. Аванессов В.С. Теоретические основы разработки знаний в тестовой форме. – М.: Изд-во Исслед. Центра проблем качества подготовки специалистов, 1995. – 95 с.
6. Комплекс нормативних документів для розробки складових системи стандартів вищої освіти. – К., 1998.
7. Колісник М. Методичне забезпечення працює на успіх. Закордонна практика викладання дисциплін // Аналітичний журнал по менеджменту «СИНЕРГІЯ». – 2003. – № 2 (6). – С. 48-53.
8. Белоус Н.В., Войтович И.В. Автоматизированная система оценивания тестовых заданий разных форм // Вестник ХНТУ. – 2006. – № 1 (24). – С. 422-426.
9. Стенли Р. Перечислительная комбинаторика: Пер. с англ. – М.: Мир, 1990. – 440 с.

Н.В. Білоус, І.В. Куцевич

Диференціальне оцінювання знань при дистанційному тестуванні

У статті досліджені основні підходи до створення тестів, наведена класифікація форм тестових завдань. У роботі авторами пропонуються принципи побудови диференціальної моделі процесу тестування, яка підтримує індивідуальність оцінювання завдань різних форм, а також наводиться технологія оцінювання ймовірності вгадування вірних відповідей.

Статья поступила в редакцию 24.06.2008.